

Large strings of consecutive smooth integers

Filip Najman

Abstract. In this note we improve an algorithm from a recent paper by Bauer and Bennett for computing a function of Erdős that measures the minimal gap size $f(k)$ in the sequence of integers at least one of whose prime factors exceeds k . This allows us to compute values of $f(k)$ for larger k and obtain new values of $f(k)$.

Keywords Pell equation, Compact representations, Smooth numbers

Mathematics subject classification (2000) 11N25, 11D09.

1 Introduction

For any integer m we let $P(m)$ be the largest prime factor of m with the convention $P(0) = P(\pm 1) = 1$. Let $\Pi_{n,k}$ be the product of k consecutive integers, starting with n , i.e.

$$\Pi_{n,k} = n(n+1) \cdots (n+k-1).$$

By a theorem of Sylvester (see [15]) $\Pi_{n,k}$ is divisible by a prime $p > k$ whenever $n > k$, and thus, following Erdős [7], we define $f(k)$ to be the least integer with the property that

$$P(\Pi_{n,f(k)}) > k,$$

if $n > k$. Standard heuristics for the size of gaps between consecutive primes lead one to expect that order of magnitude of $f(k)$ is $(\log k)^2$.

The following table gives known values of $f(k)$:

k	$f(k)$	k	$f(k)$	k	$f(k)$
1	1	13 – 40	6	61 – 113	14
2	2	41 – 46	7	114	13
3 – 4	3	47 – 58	8	115 – 150	12
5 – 12	4	59 – 60	9	151 – 178	14

The values of $f(k)$ for $k \leq 10$ were computed by Utz [16] and extended to $k \leq 42$ by Lehmer [11], to $k \leq 46$ by Ecklund and Eggleton [4], to $k \leq 73$ by Ecklund, Eggleton and Selfridge [5], [6] and finally to $k \leq 178$ by Bauer

and Bennett [1]. Bauer and Bennett in the same paper also disproved an assertion of Utz that f is monotone.

In this paper we compute the values of $f(k)$ for $k \leq 268$. Our results can be summarized in the following theorem.

Theorem 1. *For $179 \leq k \leq 268$ the values of $f(k)$ are as follows:*

k	$f(k)$
179 – 222	14
223 – 268	16

Note that Bauer and Bennett, although greatly extending the set of k such that $f(k)$ is known, by 105 values of k , did not find any values of $f(k)$ such that $f(k) > 14$. In Theorem 1 we find the new largest proven value of $f(k)$, the first after nearly 40 years (the previous being [6]).

2 The algorithm

Lemher [10] searched for two consecutive smooth integers, satisfying $P(z(z+1)) \leq p_t$, where p_t is the t -th prime. One can write $x = 2z + 1$ and see that $P(z(z+1)) \leq p_t$ iff $P(x^2 - 1) \leq p_t$. Writing $x^2 - 1 = dy^2$, where d is squarefree, leads to the Pell equation

$$x^2 - dy^2 = 1, \tag{1}$$

where $P(dy) \leq p_t$. This implies that the solutions $x + y\sqrt{d}$ we are searching for are powers of the fundamental solution $u + v\sqrt{d}$, i.e

$$x + y\sqrt{d} = (u + v\sqrt{d})^n.$$

By classic results on primitive divisors (see [3]) and the work of Lehmer (see [9]), if we want $P(y) \leq p_t$, then

$$n \leq \max \left\{ \frac{p_t + 1}{2}, 12 \right\}.$$

Thus one needs to consider only finitely many Pell equations and for each one only some of the first solutions.

After finding all pairs of smooth consecutive integers in this way, one can search through the results and find larger strings of consecutive integers.

Bauer and Bennett [1] improve on this strategy by the following clever argument: for fixed $m \leq 3$ and t suppose we are searching for all integers n satisfying

$$P(\Pi_{n,m}) \leq p_t. \tag{2}$$

We can split the indices $0, \dots, m-1$ into $\lfloor \frac{m}{4} \rfloor + \lfloor \frac{m+1}{4} \rfloor$ disjoint pairs $(i, i+2)$, where $i \equiv 0, 1 \pmod{4}$. Set $t_0 = \pi(m-1)$. By the Dirichlet principle, we can find an index i such that $(n+i)(n+i+2)$ is divisible by at most

$$N = \left\lfloor \frac{t - t_0}{\lfloor \frac{m}{4} \rfloor + \lfloor \frac{m+1}{4} \rfloor} \right\rfloor$$

of the primes from the set

$$\{p_{t_0+1}, \dots, p_t\}. \quad (3)$$

Now one writes $X = n + i + 1$ and $Y = (n + i)(n + i + 2)$, and gets

$$X^2 - DY^2 = 1,$$

where D is squarefree and divisible only by some of the first t_0 primes and at most N of the primes from the set (3). Then one proceeds exactly as Lehmer, with the difference that in this approach $P(X) \leq p_t$ also has to hold. This lowers the number of Pell equations one needs to consider from $2^t - 1$ to

$$M = -1 + 2^{t_0} \sum_{j=0}^N \binom{t - t_0}{j},$$

and equally important, reduces the size of the Pell equations.

The bottleneck of both the algorithms of Lehmer and Bauer and Bennett is solving the Pell equation. The main difficulty in solving the Pell equation is the size of the solutions, as it grows exponentially with respect to the size of the coefficient d from (1). This means that just *writing down* the solution in standard representation takes exponential time. In [12], Luca and the author used *compact representations* of the solutions to the Pell equations to overcome this difficulty and managed to extend Lehmer's results (see also [14] for another application of this approach). Note that the algorithm described in [12] is still the best up to date if one wants to find 2 or 3 consecutive smooth integers.

A compact representation of an algebraic number $\beta \in \mathbb{Q}(\sqrt{d})$ is a representation of β of the form

$$\beta = \prod_{j=1}^k \alpha_j^{2^j}, \quad (4)$$

where $\alpha_j = (a_j + b_j\sqrt{d})/c_j$, where $a_j, b_j, c_j, k \in \mathbb{Z}$, and there exist polynomial (in $\log d$) upper bounds for k , a_j , b_j and c_j . A detailed description of compact representations and their use can be found in [8]. Using compact representations, the Pell equation is solved in two steps. First, the regulator of the

appropriate quadratic field is computed. Once the regulator is known, one computes a compact representation, in polynomial time. By far the harder part of this process is the computation of the regulator. The fastest unconditional algorithms are still exponential with respect to d , while the fastest known algorithm, Buchmann's subexponential algorithm (see [2]), depends on the Generalized Riemann Hypothesis.

For our purposes, the only algorithm fast enough is Buchmann's algorithm. We will perform a check, using continued fractions, that will for each case tell us either that the output of Buchmann's algorithm is unconditionally correct or that we can disregard this case. This implies that, although we cannot unconditionally solve the Pell equation in subexponential time, we can determine unconditionally whether it has a smooth solution (and find any smooth solutions) in subexponential time.

We do this by the following method, first used in [12]: suppose we want to find the solutions of $x^2 - dy^2 = 1$, where x and y are n -smooth (d is a fixed n -smooth integer). We first compute the regulator R_d of $\mathbb{Q}(\sqrt{d})$ using Buchmann's subexponential algorithm. As mentioned, the correctness of the output R_d is dependent on the GRH, but what is unconditionally true is that the output is equal to mR_d , where m is some positive integer. From the obtained value we construct, in polynomial time, a compact representation of the m -th power of the fundamental unit, and from it a compact representation of the m -th solution $x_m + y_m\sqrt{d}$ of our Pell equation. We define z to be the n -smooth part of y_m , which is computed using the algorithm for performing modular arithmetic on compact representations from [13].

Let $x_1 + y_1\sqrt{d}$ be the fundamental solution of our Pell equation and $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^k$. If the GRH is false and Buchmann's algorithm does not return the exact value of the regulator this will mean that it might happen that we will miss some smooth solutions of our equation. Suppose this happens. This means that $m > 1$ and y_k is n -smooth for some k that is not a multiple of m . But we have $y_1|y_k$, so y_1 is also n -smooth, implying $y_1|z$ (because $y_1|y_m$). This obviously implies $y_1 \leq z$. So if something has gone wrong, we will find the solution y_1 that is smaller or equal to z and strictly smaller than y_m . We can use continued fractions to check whether such a y_1 exists. If we do not find such a y_1 (and indeed we never do, as it's existence would disprove the GRH), it means either $m = 1$ or $m > 1$ but we are not missing a smooth solution. In any case, this removes the dependence of our results on the Generalized Riemann Hypothesis. Note that although y_m does become very large, z in practice does not. This means that it is usually, for our purposes, enough to compute the first 10 or so convergents of the expansion of \sqrt{d} to a continued fraction, to prove that the hypothetical y_1 does not exist.

Using compact representation, in combination with this check, cuts down the space needed from exponential to polynomial with respect to d and the time needed to solve the Pell equation from exponential to subexponential.

In essence, in this paper we combine the clever approach of [1] and the powerful methods for finding smooth solutions of the Pell equations from [12].

It is worth noting the somewhat counter-intuitive implications of the mentioned algorithms: it is easier to solve the Pell equation (using compact representations) than it is to write down the solution in decimal notation, and it is easier to unconditionally find all smooth solutions of the Pell equation or determine that there are none (this takes subexponential time) than it is to unconditionally solve the Pell equation (this takes exponential time).

3 Results

We run our algorithm for two pairs of parameters (m, t) , these pairs being $(47, 14)$, and $(56, 16)$. Note that $p_{47} = 211$, $p_{56} = 263$ and $P(\Pi_{318,13}) = 163$, $P(\Pi_{1330,15}) = 223$, implying

$$f(k) \geq 14 \text{ for } 178 \leq k \leq 223 \text{ and } f(k) \geq 16 \text{ for } 224 \leq k \leq 268 \quad (5)$$

By proving that there is no integer $n > k$ for $178 \leq k \leq 222$ satisfying $P(\Pi_{n,14}) \leq k$ implies $f(k) = 14$ for $178 \leq k \leq 222$. In the same way one proves that $f(k) = 16$ for $223 \leq k \leq 268$.

The case $(m, t) = (47, 14)$ makes us solve

$$M_1 = -1 + 2^6 \sum_{j=0}^6 \binom{41}{j} = 342948991$$

Pell equations, while the case $(m, t) = (56, 16)$ makes us solve

$$M_2 = -1 + 2^6 \sum_{j=0}^6 \binom{50}{j} = 1168680703$$

Pell equations.

We obtain that there do not exist such integers n , thus proving Theorem 1.

Acknowledgements.

These materials are based on work financed by the National Foundation for Science, Higher Education and Technological Development of the Republic of Croatia.

References

- [1] M. Bauer and M. A. Bennett, *Prime factors of consecutive integers*, Math. Comp., **77** (2008), 2455–2459.
- [2] J. Buchmann, *A subexponential algorithm for the determination of class groups and regulators of algebraic number fields*, Séminaire de Théorie des Nombres (1990), 27–41.
- [3] R. D. Carmichael, *On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$* , Ann. of. Math. **15** (1913), 30–70.
- [4] E. F. Ecklund, Jr. and R. B. Eggleton, *Prime factors of consecutive integers*, Amer. Math. Monthly **79** (1972), 1082–1089.
- [5] E. F. Ecklund, Jr., R. B. Eggleton and J. L. Selfridge, *Factors of consecutive integers*, Proc. Man. Conference Numerical Maths., Winnipeg (1971), 155–157.
- [6] E. F. Ecklund, Jr., R. B. Eggleton and J. L. Selfridge *Consecutive integers all of whose prime factors belong to a given set*, Proc. Man. Conference Numerical Maths., Winnipeg (1971), 161–162.
- [7] P. Erdős, *On consecutive integers*, Nieuw Arch. Wisk. **3** (1955), 124–128.
- [8] M. J. Jacobson Jr., H. C. Williams, *Solving the Pell Equation*, Springer, 2009.
- [9] D. H. Lehmer, *An extended theory of Lucas functions*, Ann. Math. **31** (1930), 419–448.
- [10] D. H. Lehmer, *On a problem of Störmer*, Illinois J. Math **8** (1964), 57–79.
- [11] D. H. Lehmer, *The prime factors of consecutive integers* Amer. Math. Monthly **72** (1965), 19–20.
- [12] F. Luca and F. Najman, *On the largest prime factor of $x^2 - 1$* , Math. Comp. **80** (2011), 429–435.
- [13] F. Najman, *Compact representation of quadratic integers and integer points on some elliptic curves*, Rocky Mountain J. Math. **40** (2010), 1979–2002.
- [14] F. Najman, *Smooth values of some quadratic polynomials*, Glas. Mat. Ser III **45** (2010), 347–355.

- [15] J. J. Sylvester, *On arithmetical series*, Messenger Math. **21** (1892), 1-19, 87-120.
- [16] W. R. Utz, *A conjecture of Erdos concerning consecutive integers*, Amer. Math. Monthly **68** (1961), 896-897.

MATHEMATISCH INSTITUUT, P.O. BOX 9512, 2300 RA LEIDEN, THE NETHERLANDS

E-mail address: fnajman@math.leidenuniv.nl

AND

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ZAGREB, BIJENIČKA CESTA 30, 10000 ZAGREB, CROATIA

E-mail address: fnajman@math.hr